

STIC-Biotech/ChemLib

169457

From: Chernyshev, Olga  
Sent: Tuesday, October 25, 2005 10:14 AM  
To: STIC-Biotech/ChemLib  
Subject: 10/088,724 sequence search request

Please search SEQ ID NO: 63 in regular databases.  
Thank you very much!

Olga N. Chernyshev, Ph.D.  
AU 1649  
REM 3C89  
2-0870  
mail 4C70

RECEIVED  
OCT 25 2005  
TECH/CHEM. DIVISION  
(STIC)

\*\*\*\*\*

Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: \_\_\_\_\_  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIS: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

This page blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2005, 04:41:36 ; Search time 163 Seconds  
(without alignments)  
25.614 Million cell updates/sec

Title: US-10-088-724C-63  
Perfect score: 29  
Sequence: 1 PXXXXLTXXP 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1862994 segs, 417510619 residues  
Total number of hits satisfying chosen parameters: 1862994

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Published Applications AA.\*  
1: /cgn2\_6/prodata/2/pubppaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/prodata/2/pubppaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/2/pubppaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/2/pubppaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/2/pubppaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/2/pubppaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/prodata/2/pubppaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/prodata/2/pubppaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/prodata/2/pubppaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/prodata/2/pubppaa/US09C\_PUBCOMB.pep.\*  
11: /cgn2\_6/prodata/2/pubppaa/US09C\_NEW\_PUB.pep.\*  
12: /cgn2\_6/prodata/2/pubppaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/prodata/2/pubppaa/US10C\_PUBCOMB.pep.\*  
14: /cgn2\_6/prodata/2/pubppaa/US10C\_PUBCOMB.pep.\*  
15: /cgn2\_6/prodata/2/pubppaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/prodata/2/pubppaa/US10C\_PUBCOMB.pep.\*  
17: /cgn2\_6/prodata/2/pubppaa/US10C\_PUBCOMB.pep.\*  
18: /cgn2\_6/prodata/2/pubppaa/US10C\_PUBCOMB.pep.\*  
19: /cgn2\_6/prodata/2/pubppaa/US11\_PUBCOMB.pep.\*  
20: /cgn2\_6/prodata/2/pubppaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/prodata/2/pubppaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/prodata/2/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	75.9	83	US-10-437-963-153122	Sequence 153122,
2	22	75.9	103	US-10-425-115-237142	Sequence 237142,
3	22	75.9	122	US-10-424-599-236119	Sequence 236119,
4	22	75.9	127	US-10-767-701-44795	Sequence 44795, A
5	22	75.9	154	US-10-856-499-775	Sequence 775, App
6	22	75.9	215	US-10-108-260A-3211	Sequence 3211, App
7	22	75.9	286	US-10-856-499-920	Sequence 920, App
8	22	75.9	306	US-10-437-963-172644	Sequence 172644,
9	22	75.9	351	US-11-097-143-38586	Sequence 38586, A
10	22	75.9	381	US-10-128-714-3245	Sequence 3245, Ap
11	22	75.9	390	US-10-437-963-139856	Sequence 139856,

12	22	75.9	405	14	US-10-128-714-8245	Sequence 8245, Ap
13	22	75.9	425	16	US-10-343-663A-22	Sequence 22, Appl
14	22	75.9	425	16	US-10-343-663A-53	Sequence 53, Appl
15	22	75.9	495	16	US-10-437-963-113675	Sequence 113675,
16	22	75.9	580	16	US-10-723-860-2696	Sequence 2696, Ap
17	22	75.9	1051	16	US-10-739-930-5685	Sequence 5685, Ap
18	22	75.9	19652	15	US-10-084-846A-7	Sequence 7, Appl1
19	22	72.4	10	9	US-09-801-784-36	Sequence 36, Appl
20	21	72.4	10	16	US-10-754-642-36	Sequence 36, Appl
21	21	72.4	22	18	US-10-862-195-706	Sequence 706, App
22	21	72.4	36	16	US-10-754-642-1	Sequence 1, Appl1
23	21	72.4	36	17	US-10-864-803-1	Sequence 37, Appl
24	21	72.4	37	9	US-09-801-784-37	Sequence 37, Appl
25	21	72.4	37	16	US-10-754-642-37	Sequence 3, Appl1
26	21	72.4	37	17	US-10-864-803-3	Sequence 40043, A
27	21	72.4	44	9	US-09-864-761-40043	Sequence 268687,
28	21	72.4	55	16	US-10-425-115-268687	Sequence 57, Appl
29	21	72.4	57	15	US-10-182-243-57	Sequence 144749,
30	21	72.4	57	16	US-10-437-963-144749	Sequence 2764, Ap
31	21	72.4	65	11	US-09-864-408A-2754	Sequence 149291,
32	21	72.4	68	15	US-10-424-599-149291	Sequence 236598,
33	21	72.4	74	16	US-10-425-115-236598	Sequence 27928,
34	21	72.4	87	15	US-10-424-599-27928	Sequence 324968,
35	21	72.4	90	16	US-10-425-115-324968	Sequence 484, App
36	21	72.4	94	18	US-10-972-024-484	Sequence 173090,
37	21	72.4	99	15	US-10-424-599-173090	Sequence 182814,
38	21	72.4	100	15	US-10-424-599-182814	Sequence 357230,
39	21	72.4	108	16	US-10-425-115-357230	Sequence 279805,
40	21	72.4	109	15	US-10-424-599-279805	Sequence 254115,
41	21	72.4	109	16	US-10-425-115-254115	Sequence 247418,
42	21	72.4	113	16	US-10-425-115-247418	Sequence 275463,
43	21	72.4	117	15	US-10-424-599-275463	Sequence 161257,
44	21	72.4	123	16	US-10-437-963-161257	Sequence 2137, Ap
45	21	72.4	129	15	US-10-094-749-2137	

## ALIGNMENTS

RESULT 1  
US-10-437-963-153122  
Sequence 153122, Application US/10437963  
Publication No. US20040123343A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovacic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barzak, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437, 963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 153122  
LENGTH: 83  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_53105C.1.pep  
US-10-437-963-153122

Query Match 75.9%; Score 22; DB 16; Length 83;  
Best Local Similarity 40.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Cy 1 PXXXXLTXXP 10  
Db 55 PSSAELTSAP 64

RESULT 2  
US-10-425-115-237142  
; Sequence 237142, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 237142  
; LENGTH: 103  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_147860C.1.pep  
US-10-425-115-237142  
Query Match  
Best Local Similarity 75.9%; Score 22; DB 16; Length 103;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 PXXXXLTXXP 10  
DB 76 PASTSLTSP 85  
RESULT 3  
US-10-424-599-236119  
; Sequence 236119, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 236119  
; LENGTH: 122  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(122)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_55243C.1.pep  
US-10-424-599-236119  
Query Match  
Best Local Similarity 75.9%; Score 22; DB 15; Length 122;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 PXXXXLTXXP 10  
DB 26 PSAGALTTTP 35  
RESULT 4  
US-10-767-701-44795  
; Sequence 44795, Application US/10767701

; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 44795  
; LENGTH: 127  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C4945\_2.pep  
US-10-767-701-44795  
Query Match  
Best Local Similarity 75.9%; Score 22; DB 16; Length 127;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 PXXXXLTXXP 10  
DB 32 PLSSSLTSP 41  
RESULT 5  
US-10-856-499-775  
; Sequence 775, Application US/10856499  
; Publication No. US20040259145A1  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Marion  
; APPLICANT: Shenk, Michael A.  
; APPLICANT: McGrath, Annette  
; APPLICANT: Glenn, Matthew  
; TITLE OF INVENTION: Compositions and Methods for the  
; FILE REFERENCE: 11000.1021C2  
; CURRENT APPLICATION NUMBER: US/10/856,499  
; CURRENT FILING DATE: 2004-05-28  
; NUMBER OF SEQ ID NOS: 2370  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 775  
; LENGTH: 154  
; TYPE: PRT  
; ORGANISM: Eucalyptus grandis  
US-10-856-499-775  
Query Match  
Best Local Similarity 75.9%; Score 22; DB 16; Length 154;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 PXXXXLTXXP 10  
DB 140 PSSSLTSP 149  
RESULT 6  
US-10-108-260A-3211  
; Sequence 3211, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1 full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3211  
; LENGTH: 215

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2005, 04:38:06 ; Search time 41 Seconds  
(without alignments)  
18.207 Million cell updates/sec

Title: US-10-088-724c-63

Perfect score: 29  
Sequence: 1 PXXXXLTXXP 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued Patents AA:  
1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/1/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/1aa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	75.9	128	US-09-248-796A-22571	Sequence 22571, A
2	22	75.9	154	US-09-640-211A-775	Sequence 775, App
3	22	75.9	221	US-09-252-991A-26404	Sequence 26404, A
4	22	75.9	286	US-09-640-211A-920	Sequence 920, App
5	21	72.4	36	US-08-460-617-2	Sequence 2, Appli
6	21	72.4	36	PCT-US96-08730-1	Sequence 1, Appli
7	21	72.4	36	PCT-US96-08730-2	Sequence 2, Appli
8	21	72.4	37	US-08-460-617-3	Sequence 3, Appli
9	21	72.4	37	US-08-460-617-4	Sequence 4, Appli
10	21	72.4	37	PCT-US96-08730-15	Sequence 15, Appli
11	21	72.4	38	US-09-428-498-1	Sequence 1, Appli
12	21	72.4	53	US-09-513-999C-6581	Sequence 6581, Ap
13	21	72.4	134	US-09-644-460-33	Sequence 33, Appli
14	21	72.4	148	US-09-513-999C-5275	Sequence 5275, Ap
15	21	72.4	174	US-09-252-991A-26812	Sequence 26812, A
16	21	72.4	387	US-09-916-109-6	Sequence 6, Appli
17	21	72.4	387	US-10-211-412B-6	Sequence 6, Appli
18	21	72.4	388	US-08-894-772-2	Sequence 2, Appli
19	21	72.4	388	US-09-207-844-2	Sequence 2, Appli
20	21	72.4	469	US-09-134-000C-5065	Sequence 5065, Ap
21	21	72.4	483	US-09-916-109-4	Sequence 4, Appli
22	21	72.4	483	US-10-211-412B-4	Sequence 4, Appli
23	21	72.4	483	US-09-538-092-1162	Sequence 1162, Ap
24	21	72.4	508	US-09-252-991A-18823	Sequence 18823, A
25	21	72.4	552	US-09-385-028-3	Sequence 3, Appli
26	21	72.4	552	US-09-726-614-3	Sequence 3, Appli
27	21	72.4	552	US-09-385-040-3	Sequence 3, Appli

28	21	72.4	554	4	US-09-489-039A-12959	Sequence 12959, A
29	21	72.4	568	4	US-09-252-991A-19968	Sequence 19968, A
30	21	72.4	632	1	US-08-443-104-4	Sequence 4, Appli
31	21	72.4	632	1	US-08-442-859-4	Sequence 4, Appli
32	21	72.4	632	2	US-08-398-489-4	Sequence 4, Appli
33	21	72.4	632	5	PCT-US95-05534-4	Sequence 4, Appli
34	21	72.4	752	1	US-08-420-235B-21	Sequence 21, Appli
35	21	72.4	752	3	US-08-793-624-21	Sequence 21, Appli
36	21	72.4	752	5	PCT-US95-10194-21	Sequence 21, Appli
37	21	72.4	1043	4	US-09-252-991A-28885	Sequence 28885, A
38	21	72.4	1222	2	US-08-682-517-15	Sequence 15, Appli
39	21	72.4	1252	2	US-08-682-517-9	Sequence 9, Appli
40	20	69.0	19	2	US-08-613-235-6	Sequence 6, Appli
41	20	69.0	37	2	US-08-460-617-5	Sequence 5, Appli
42	20	69.0	37	2	US-08-460-617-6	Sequence 6, Appli
43	20	69.0	37	5	PCT-US96-08730-4	Sequence 4, Appli
44	20	69.0	37	5	PCT-US96-08730-5	Sequence 5, Appli
45	20	69.0	61	1	US-08-112-208C-5	Sequence 5, Appli

## ALIGNMENTS

```

RESULT 1
US-09-248-796A-22571
; Sequence 22571, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248, 796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074, 725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096, 409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 22571
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-22571

Query Match      75.9%; Score 22; DB 4; Length 128;
Best Local Similarity 40.0%; Pred. No. 91;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 PXXXXLTXXP 10
Db      77 PFTSTLTSDP 86
      || |
      || |

RESULT 2
US-09-640-211A-775
; Sequence 775, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021CIU
; CURRENT APPLICATION NUMBER: US/09/640, 211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 775
; LENGTH: 154
; TYPE: PRT

```

ORGANISM: Euca1yptus grandis  
US-09-640-211A-775

Query Match 75.9%; Score 22; DB 4; Length 154;  
Best Local Similarity 40.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 PXXXXLTXXP 10  
DB 140 PSSSLTSP 149

RESULT 3  
US-09-252-991A-26404

Sequence 26404, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT FILING DATE: 1999-02-18

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 26404

LENGTH: 221

TYPE: PR

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-26404

Query Match 75.9%; Score 22; DB 4; Length 221;  
Best Local Similarity 40.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 PXXXXLTXXP 10  
DB 64 PLASTTSP 73

RESULT 4  
US-09-640-211A-920

Sequence 920, Application US/09640211A

Patent No. 6833446

GENERAL INFORMATION:

APPLICANT: Wood, Marion

APPLICANT: Shenk, Michael A.

APPLICANT: McGrath, Annette

APPLICANT: Glenn, Matthew

TITLE OF INVENTION: Compositions and Methods for the

FILE REFERENCE: 11000.1021C1U

CURRENT APPLICATION NUMBER: US/09/640,211A

CURRENT FILING DATE: 2000-08-16

NUMBER OF SEQ ID NOS: 2368

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 920

LENGTH: 286

TYPE: PR

ORGANISM: Euca1yptus grandis

US-09-640-211A-920

Query Match 75.9%; Score 22; DB 4; Length 286;  
Best Local Similarity 40.0%; Pred. No. 2e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 PXXXXLTXXP 10  
DB 273 PSSSLTSP 282

RESULT 5  
US-08-460-617-2

Sequence 2, Application US/08460617

Patent No. 5914114

GENERAL INFORMATION:

APPLICANT: Cassels, Frederick J

TITLE OF INVENTION: Method of Raising Antibodies Against E.

TITLE OF INVENTION: Coli

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESS: Hendricks and Associates

STREET: 9669 A Main Street

CITY: Fairfax

STATE: VA

COUNTRY: US

ZIP: 22031

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/460,617

FILING DATE: 02-JUN-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Hendricks, Glenna

REGISTRATION NUMBER: 32,535

REFERENCE/DOCKET NUMBER: cas460

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-425-4250

TELEFAX: 703-425-2767

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 36 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

ORGANISM: E coli

US-08-460-617-2

Query Match 72.4%; Score 21; DB 2; Length 36;  
Best Local Similarity 40.0%; Pred. No. 49;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 PXXXXLTXXP 10  
DB 26 PASVALTSP 35

RESULT 6  
PCT-US96-08730-1

Sequence 1, Application PC/TUS9608730

GENERAL INFORMATION:

APPLICANT: Cassels, Frederick

APPLICANT: Anderson, Jeffrey

APPLICANT: Carter, John Mark

TITLE OF INVENTION: Methods of Raising Antibodies Against E.

TITLE OF INVENTION: Coli of the Family CSF-CFA./1

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESS: Glenna Hendricks

STREET: P.O. Box 2509

CITY: Fairfax

STATE: VA

COUNTRY: USA

ZIP: 22031

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2005, 04:36:30 ; Search time 176 Seconds  
(without alignments)  
29.095 Million cell updates/sec

Title: US-10-088-724c-63  
Perfect score: 29  
Sequence: 1 PXXXXLTXXP 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues  
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	79.3	492	2	Q6CP98
2	23	79.3	799	2	Q7ZGR4
3	23	79.3	942	2	Q6PZ6
4	23	79.3	1320	1	TCOF MOUSE
5	23	79.3	1320	2	Q7TFZ
6	22	75.9	170	2	Q71EB4
7	22	75.9	215	1	RM43 HUMAN
8	22	75.9	286	2	Q6DGH5
9	22	75.9	301	2	Q7PRR4
10	22	75.9	306	2	Q7Q3C2
11	22	75.9	306	2	Q7XN0
12	22	75.9	307	2	Q9NTQ6
13	22	75.9	315	2	Q7QZ29
14	22	75.9	319	2	Q9VFS5
15	22	75.9	377	2	Q9VOY0
16	22	75.9	387	2	Q81EP1
17	22	75.9	390	2	Q9ASU9
18	22	75.9	425	1	PLA2 MOUSE
19	22	75.9	462	1	MB12 ARATH
20	22	75.9	462	2	Q6M187
21	22	75.9	462	2	Q6M187
22	22	75.9	559	2	Q6M187
23	22	75.9	589	1	STFB SCHPO
24	22	75.9	720	2	Q68A49
25	22	75.9	725	2	Q9CV93
26	22	75.9	749	2	Q967D9
27	22	75.9	762	2	Q7SD72
28	22	75.9	847	1	RS62 RAT
29	22	75.9	902	2	Q81Q17
30	22	75.9	903	2	Q967D8
31	22	75.9	903	2	Q9VQY1

32	22	75.9	970	2	Q867Z4	Q867Z4 drosophila
33	22	75.9	1051	2	Q9SE97	Q9SE97 arabidopsis
34	22	75.9	1072	2	Q6FKH4	Q6FKH4 candida gla
35	22	75.9	1237	2	Q7SGA0	Q7SGA0 neurospora
36	22	75.9	1245	2	Q8XIE3	Q8XIE3 raltstonia s
37	22	75.9	1508	2	Q6NR34	Q6NR34 drosophila
38	22	75.9	1508	2	Q9VQY2	Q9VQY2 drosophila
39	22	75.9	1531	2	Q9VQY2	Q9VQY2 drosophila
40	22	75.9	1625	2	Q7ND10	Q7ND10 gloeobacter
41	22	75.9	4210	2	Q89J13	Q89J13 bradyrhizob
42	22	75.9	4782	2	Q8K1G6	Q8K1G6 mus musculu
43	21	72.4	59	2	Q54136	Q54136 saccharopol
44	21	72.4	79	2	Q7ULJ8	Q7ULJ8 rhodospirell
45	21	72.4	111	2	Q9Y9B2	Q9Y9B2 aeropyrum p

## ALIGNMENTS

RESULT 1	ID	PRELIMINARY	PRT	492 AA.
Q6CP98	Q6CP98	PRELIMINARY	PRT	492 AA.
AC	Q6CP98			
DT	25-OCT-2004 (T-EMBLrel. 28, Created)			
DT	25-OCT-2004 (T-EMBLrel. 28, Last sequence update)			
DT	25-OCT-2004 (T-EMBLrel. 28, Last annotation update)			
DE	Kluyveromyces lactis strain NRRL Y-1140 chromosome E of strain NRRL Y-1140 of Kluyveromyces lactis.			
GN	ORFNames=KLA0E064795;			
OC	Kluyveromyces lactis NRRL Y-1140.			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Kluyveromyces.			
OX	NCBI_TaxID=284590;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NRRL Y-1140;			
RG	Genolevures:			
RA	Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S., LaFontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E., Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V., Barney S., Blanchin S., Beckerlich J.M., Beyne E., Bleykasten C., Boistrane A., Boyer J., Cattolico L., Confanioli F., de Daruvar A., Despons L., Fabre B., Fairhead C., Ferry-Dumazet H., Groppi A., Hantaye F., Hennequin C., Janniaux N., Joyet P., Kachouri R., Karrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H., Nicaud J.M., Nikolski M., Ozias S., Ozier-Kalogeropoulos O., Pellenz S., Portier S., Richard G.F., Straub M.L., Suleau A., Swemene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B., Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A., Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J., Wincker P., Souciet J.L., "Genome evolution in yeasts."			
RL	Nature 430:35-44 (2004).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NRRL Y-1140;			
RG	Genoscope;			
RA	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.			
RL	-1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.			
DR	EMBL: CR82125; CAG9328.1; -			
DR	InterPro: IPR001092; HLH_basic.			
DR	Pfam: PF00010; HLH.1.			
DR	SMART: SM00353; HLH.1.			
DR	PROSITE: PS50888; HLH.1.			
SQ	SEQUENCE 492 AA; 55864 MW; 0C8879277AB7DAE0 CRC64;			
Query Match	79.3%;	Score 23;	DB 2;	Length 492;
Best local Similarity	40.0%;	Pred. No. 2.8e+02;		
Matches	4;	Conservative	0;	Mismatches 6;
			Indels	0;
			Gaps	0;
OY	1 PXXXXLTXXP 10			
DB	239 FTSSSLTSTP 248			

```

RESULT 2
ID Q7ZGR4 PRELIMINARY; PRT; 799 AA.
AC Q6PFZ6;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxId=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22347655; PubMed=12459786; DOI=10.1038/nature01200;
RA Allen T.M., Altfield M., Yu X.G., Johnston M.N., Agrawal D.,
RA Korber B.T., Montefiori D.C., O'Connor D.H., Davis B.T., Lee P.K.,
RA Maier E.L., Harlow J., Goulder P.J.R., Brander C., Rosenberg E.S.,
RA Walker B.D.;
RT "HIV-1 superinfection despite broad CD8+ T-cell responses containing
RT replication of the primary virus."
RN Nature 420:434-439(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Allen T.M., Altfield M., Yu X.G., Johnston M.N., Agrawal D.,
RA Korber B.T., Montefiori D.C., O'Connor D.H., Davis B.T., Lee P.K.,
RA Maier E.L., Harlow J., Goulder P.J.R., Brander C., Rosenberg E.S.,
RA Walker B.D.;
RT Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY247251; AAP12634.1; -.
DR HSSP; P05877; IN10.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0015028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Polypotein;
DR Transmembrane.
FT NON TER 799
SQ SEQUENCE 799 AA; 90679 MW; 655B25158B31460A CRC64;

Query Match 79.3%; Score 23; DB 2; Length 799;
Best Local Similarity 40.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 PXXXXLTXXP 10
DB 152 PTTSLTTP 161

RESULT 3
ID Q6PFZ6 PRELIMINARY; PRT; 942 AA.
AC Q6PFZ6;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Tcof1 protein.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

```

```

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko U., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Locuelli N.A., Peters G.J., Abramson R.D., Miliady S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RT Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC057342; AAH57342.1; -.
DR InterPro; IPR006594; L1SH.
DR InterPro; IPR003993; Treacle.
DR Pfam; PF03546; Treacle; 1.
DR PRINTS; PR01503; TREACLE.
DR SMART; SM00667; L1SH; 1.
DR PROSITE; PS50896; L1SH; 1.
SQ SEQUENCE 942 AA; 94643 MW; 861CB009A920DAB CRC64;

Query Match 79.3%; Score 23; DB 2; Length 942;
Best Local Similarity 40.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 PXXXXLTXXP 10
DB 425 PAAATLTSP 434

RESULT 4
ID TCOF MOUSE STANDARD; PRT; 1320 AA.
AC 008784; 008857;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Treacle protein (Treacher Collins syndrome protein homolog).
GN Name=Tcof1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;
RX MEDLINE=97445113; PubMed=9299440; DOI=10.1006/dbrc.1997.7229;
RA Paznekas W.A., Zhang N., Gridley T., Jabs E.W.;
RT "Mouse TCOF1 is expressed widely, has motifs conserved in nuclear
RT phosphoproteins, and maps to chromosome 18."
RL Biochem. Biophys. Res. Commun. 238:1-6(1997).
RN [2]
RP SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE.
RX MEDLINE=97301769; PubMed=9158147; DOI=10.1093/hmg/6.5.727;
RA Dixon J., Hovanes K., Shiang R., Dixon M.J.;
RT "Sequence analysis, identification of evolutionary conserved motifs
RT and expression analysis of murine tcof1 provide further evidence for a
RT potential function for the gene and its human homologue, TCOF1."
RN Hum. Mol. Genet. 6:727-737(1997).
RN [3]
RP SEQUENCE OF 1-1314 FROM N.A.
RX STRAIN=C57BL/6; TISSUE=Brain;

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 26, 2005, 04:37:41 ; Search time 39 seconds  
(without alignments)  
24.671 Million cell updates/sec

Title: US-10-088-724C-63

Perfect score: 29

Sequence: 1 PXXXXLTXXP 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database:

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79.3	1320	2	JC5630	TCOP1 protein - mo
2	75.9	462	2	TS2115	myosinase-binding
3	75.9	462	2	B96560	hypothetical prote
4	75.9	580	2	T43481	probable mucin DKF
5	75.9	589	2	T38232	phenylalanine-CPNA
6	75.9	847	2	A56039	Gpase-activating
7	72.4	111	2	D72466	hypothetical prote
8	72.4	262	2	T33597	hypothetical prote
9	72.4	276	2	T49123	hypothetical prote
10	72.4	298	2	S69241	Dreg-5 protein - f
11	72.4	432	2	T20536	hypothetical prote
12	72.4	459	2	A11874	two-component sens
13	72.4	484	2	T16595	hypothetical prote
14	72.4	501	2	D96551	hypothetical prote
15	72.4	506	2	A54190	cerebroside-sulfat
16	72.4	546	2	AG0586	phosphoglucomutase
17	72.4	546	2	G64803	phosphoglucomutase
18	72.4	546	2	G85568	phosphoglucomutase
19	72.4	546	2	G90718	phosphoglucomutase
20	72.4	547	2	E70650	phosphoglucomutase
21	72.4	578	2	B40046	tetracycline resis
22	72.4	614	2	T10862	phaseolin G-box b1
23	72.4	637	2	T20981	hypothetical prote
24	72.4	642	2	T10861	phaseolin G-box b1
25	72.4	925	2	T00334	hypothetical prote
26	72.4	1010	2	T31367	lola-like protein
27	72.4	1045	2	JC5795	CDP protein - hum
28	72.4	1103	2	T42022	probable chitin sy
29	72.4	1128	1	T08322	plasmid replicatio

30	72.4	1176	2	A49848	nitrite reductase
31	72.4	1560	2	T42727	proliferation pote
32	72.4	1840	2	T30250	GTL protein - mous
33	72.4	2100	2	T03223	probable polyketid
34	69.0	102	2	B72652	hypothetical prote
35	69.0	144	1	RMWTV	T-cell receptor be
36	69.0	145	2	A13068	hypothetical prote
37	69.0	145	2	B98218	hypothetical prote
38	69.0	153	2	S75561	hypothetical prote
39	69.0	171	2	A41467	limbal protein c
40	69.0	192	2	F71169	hypothetical prote
41	69.0	218	2	B47538	bcl-2-associated p
42	69.0	222	2	E75498	conserved hypothet
43	69.0	234	2	G83698	hypothetical prote
44	69.0	237	2	G75261	cytidylate kinase
45	69.0	238	1	Q0BE74	US4 protein - huma

#### ALIGNMENTS

RESULT 1  
JC5630  
TCOP1 protein - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 14-Oct-1997 #sequence\_revision 07-Nov-1997 #text\_change 09-Jul-2004  
C/Accession: JC5630  
R/Paznekas, W.A.; Zhang, N.; Gridley, T.; Jabs, E.W.  
Biochem. Biophys. Res. Commun. 238, 1-6, 1997  
A/Title: Mouse ~~TCOP1~~ ~~is expressed widely~~ ~~has motifs~~ conserved in nucleolar phosphoprotei  
A/Reference number: JC5630; PMID:97445113; PMID:9299440  
A/Accession: JC5630  
A/Status: nucleic acid sequence not shown  
A/Molecule type: mRNA  
A/Residues: 1-1320 <PAZ>  
A/Cross-references: UNIPROT:O08784; DDBJ:AF001794; NID:G2109458; PIDN:AB71347.1; PID:G21  
C/Comment: This protein is a nucleolar phosphoprotein with 82 potential phosphorylation s  
ing.

Query Match 79.3%; Score 23; DB 2; Length 1320;  
Best Local Similarity 40.0%; Pred. No. 1.3e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 PXXXXLTXXP 10  
Db 425 PAAATLTSP 434

RESULT 2  
TS2115  
myosinase-binding protein [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 09-Jul-2004  
C/Accession: TS2115  
R/Benedetti, C.E.; Turcineelli, S.R.; Capella, A.N.; Arruda, P.  
submitted to the EMBL Data Library, March 1998  
A/Description: Isolation of an Arabidopsis cDNA specifically expressed in flowers and hon

A/Reference number: Z25962  
A/Accession: TS2115  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-462 <BEN>  
A/Cross-references: UNIPROT:O65187; EMBL:AF054906; PIDN:AAC08601.1

Query Match 75.9%; Score 22; DB 2; Length 462;  
Best Local Similarity 40.0%; Pred. No. 87;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 PXXXXLTXXP 10  
Db 148 PISSLTTP 157

RESULT 3  
B96560  
hypothetical protein F5F19.10 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C/Accession: B96560  
R/Title: Logis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Hutzar, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Maritali, R.; Rooney, M.; Rowley, D.; Rowley, D.; Sekano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A66141; MUID:21016719; PMID:11130712  
A/Accession: B96560  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-462 <STO>  
A/Cross-references: UNIPROT:Q9SAV0; GB:AE005173; NID:g4220451; PIDN:AAD12678.1; GSPDB:GN  
C/Genetics:  
A/Map position: 1  
A/Gene: F5F19.10

Query Match 75.9%; Score 22; DB 2; Length 462;  
Best Local Similarity 40.0%; Pred. No. 87;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 PXXXXLTXXP 10  
DB 148 PISSSLTTP 157

RESULT 4  
T43481

probable mucin DKFZp434C196.1 - human (fragment)  
N/Alternate names: Protein DKFZp434B0635.1

C/Species: Homo sapiens (man)  
C/Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
C/Accession: T43481; T34549; T17264

R/Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, December 1999

A/Reference number: Z22514

A/Accession: T43481  
A/Molecule type: mRNA

A/Residues: 1-580 <AAA>  
A/Cross-references: UNIPROT:Q9UF83; EMBL:AL133561; NID:g6599133; PIDN:CAB63715.1; PID:g6

R/Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, October 1999

A/Reference number: Z21540

A/Accession: T34549  
A/Molecule type: mRNA

A/Residues: 262-580 <POU1>  
A/Cross-references: EMBL:AL122069; NID:g6102864; PIDN:CAB59245.2; PID:g7018420

R/Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, September 1999

A/Reference number: Z18723

A/Accession: T17264  
A/Molecule type: mRNA

A/Residues: 262-580 <POU2>  
A/Cross-references: EMBL:AL117481; NID:g5911956; PIDN:CAB5954.1; PID:g5911959

A/Experimental source: adult testis; clone DKFZp434B061  
C/Genetics:

A/Note: DKFZp434C196.1; DKFZp434B063.1; DKFZp434B061.1  
Query Match 75.9%; Score 22; DB 2; Length 580;  
Best Local Similarity 40.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 PXXXXLTXXP 10  
DB 20 PSTASLTTP 29

RESULT 5  
T38232

phenylalanine-tRNA ligase (EC 6.1.1.20) alpha chain [similarity] - fission yeast (Schizosaccharomyces pombe)  
C/Species: Schizosaccharomyces pombe  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C/Accession: T38232

R/Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, February 1998

A/Reference number: Z21780

A/Accession: T38232  
A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA  
A/Residues: 1-589 <MUR>

A/Cross-references: UNIPROT:Q42849; EMBL:AL021813; PIDN:CAA16986.1; GSPDB:GN00066; SPDB:G  
A/Experimental source: strain 972h(-); cosmid c23A1

C/Genetics:  
A/Gene: SPDB:SPAC23A1.12c

A/Map position: 1  
A/Intons: 20/1

C/Superfamily: Yeast cytosolic phenylalanine-tRNA ligase alpha chain  
C/Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 75.9%; Score 22; DB 2; Length 589;  
Best Local Similarity 40.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 PXXXXLTXXP 10  
DB 90 PSTSLTISP 99

RESULT 6  
A56039

GTPase-activating protein Gap1(m) - rat  
C/Species: Rattus norvegicus (Norway rat)  
C/Date: 01-Dec-1995 #sequence\_revision 01-Dec-1995 #text\_change 09-Jul-2004  
C/Accession: A56039

R/Maekawa, M.; Li, S.; Iwamatsu, A.; Morishita, T.; Yokota, K.; Inai, Y.; Kohsaka, S.; N  
Mol. Cell. Biol. 14, 6879-6885, 1994

A/Title: A novel mammalian Ras GTPase-activating protein which has phospholipid-binding  
A/Reference number: A56039; MUID:95021216; PMID:7935405

A/Accession: A56039  
A/Status: preliminary; nucleic acid sequence not shown

A/Molecule type: mRNA  
A/Residues: 1-847 <MA>

A/Cross-references: UNIPROT:Q63713; GB:D30734; NID:g559374; PIDN:BA06398.1; PID:dt00696;  
A/Experimental source: brain

F/355-567/Domain: ras-specific GAP catalytic domain homology <GAP>  
F/602-702/Domain: pleckstrin repeat homology <PLK>

Query Match 75.9%; Score 22; DB 2; Length 847;  
Best Local Similarity 40.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 PXXXXLTXXP 10  
DB 6 PAAALTEAP 15

RESULT 7  
D72466

hypothetical protein APE2375 - Aeropyrum pernix (strain K1)  
C/Species: Aeropyrum pernix  
C/Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C/Accession: D72466

R/Tawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatakeyama, Y.; Jin-no, K.; Takah  
awa, H.; Takamiya, M.; Maeda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
DNA Res. 6, 83-101, 1999